

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: FROMMER, Wolf-Bernd
- (ii) TITLE OF INVENTION: DNA SEQUENCES FOR AN AMINO ACID TRANSPORTER, PLASMIDS, BACTERIA, YEASTS AND PLANTS CONTAINING A TRANSPORTER AND THEIR USE
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
 - (B) STREET: 1180 Avenue of the Americas
 - (C) CITY: New York
 - (D) STATE: NY
 - (E) COUNTRY: US
 - (F) ZIP: 10036-8403
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/362,512
 - (B) FILING DATE: 05-JAN-1995
 - (C) CLASSIFICATION: 800
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/007,636
 - (B) FILING DATE: 21-JAN-1993
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Meilman, Edward A.
 - (B) REGISTRATION NUMBER: 24,735
 - (C) REFERENCE/DOCKET NUMBER: P/951-107
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 382-0700
 - (B) TELEFAX: (212) 382-0888
 - (C) TELEX: 236925

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1685 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Arabidopsis thaliano

0985474-051404

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 57..1511

(D) OTHER INFORMATION: /note= "amino acid transporter"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTTAAAACAT TTATTTTATC TTCTTCTGTG TCTCTCTTTC TCTTTCTCTC ATCACT	56
ATG AAG AGT TTC AAC ACA GAA GGA CAC AAC CAC TCC ACG GCG GAA TCC	104
Met Lys Ser Phe Asn Thr Glu Gly His Asn His Ser Thr Ala Glu Ser	
1 5 10 15	
GGC GAT GCC TAC ACC GTG TCG GAC CCG ACA AAG AAC GTC GAT GAA GAT	152
Gly Asp Ala Tyr Thr Val Ser Asp Pro Thr Lys Asn Val Asp Glu Asp	
20 25 30	
GGT CGA GAG AAG CGT ACC GGG ACG TGG CTT ACG GCG AGT GCG CAT ATT	200
Gly Arg Glu Lys Arg Thr Gly Thr Trp Leu Thr Ala Ser Ala His Ile	
35 40 45	
ATC ACG GCG GTG ATA GGC TCC GGA GTG TTG TCT TTA GCA TGG GCT ATA	248
Ile Thr Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile	
50 55 60	
GCT CAG CTT GGT TGG ATC GCA GGG ACA TCG ATC TTA CTC ATT TTC TCG	296
Ala Gln Leu Gly Trp Ile Ala Gly Thr Ser Ile Leu Leu Ile Phe Ser	
65 70 75 80	
TTC ATT ACT TAC TTC ACC TCC ACC ATG CTT GCC GAT TGC TAC CGT GCG	344
Phe Ile Thr Tyr Phe Thr Ser Thr Met Leu Ala Asp Cys Tyr Arg Ala	
85 90 95	
CCG GAT CCC GTC ACC GGA AAA CGG AAT TAC ACT TAC ATG GAC GTT GTT	392
Pro Asp Pro Val Thr Gly Lys Arg Asn Tyr Thr Tyr Met Asp Val Val	
100 105 110	
CGA TCT TAC CTC GGT GGT AGG AAA GTG CAG CTC TGT GGA GTG GCA CAA	440
Arg Ser Tyr Leu Gly Gly Arg Lys Val Gln Leu Cys Gly Val Ala Gln	
115 120 125	
TAT GGG AAT CTG ATT GGG GTC ACT GTT GGT TAC ACC ATC ACT GCT TCT	488
Tyr Gly Asn Leu Ile Gly Val Thr Val Gly Tyr Thr Ile Thr Ala Ser	
130 135 140	
ATT AGT TTG GTA GCG GTA GGG AAA TCG AAC TGC TTC CAC GAT AAA GGG	536
Ile Ser Leu Val Ala Val Gly Lys Ser Asn Cys Phe His Asp Lys Gly	
145 150 155 160	
CAC ACT GCG GAT TGT ACT ATA TCG AAT TAT CCG TAT ATG GCG GTT TTT	584
His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe	
165 170 175	
GGT ATC ATT CAA GTT ATT CTT AGC CAG ATC CCA AAT TTC CAC AAG CTC	632
Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu	
180 185 190	
TCT TTT CTT TCC ATT ATG GCC GCA GTC ATG TCC TTT ACT TAT GCA ACT	680
Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr	

0985474-051401

195	200	205	
ATT GGA ATC GGT CTA GCC ATC GCA ACC GTC GCA GGT GGG AAA GTG GGT Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly 210 215 220			728
AAG ACG AGT ATG ACG GGC ACA GCG GTT GGA GTA GAT GTA ACC GCA GCT Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala 225 230 235 240			776
CAA AAG ATA TGG AGA TCG TTT CAA GCG GTT GGG GAC ATA GCG TTC GCC Gln Lys Ile Trp Arg Ser Phe Gln Ala Val Gly Asp Ile Ala Phe Ala 245 250 255			824
TAT GCT TAT GCC ACG GTT CTC ATC GAG ATT CAG GAT ACA CTA AGA TCT Tyr Ala Tyr Ala Thr Val Leu Ile Glu Ile Gln Asp Thr Leu Arg Ser 260 265 270			872
AGC CCA GCT GAG AAC AAA GCC ATG AAA AGA GCA AGT CTT GTG GGA GTA Ser Pro Ala Glu Asn Lys Ala Met Lys Arg Ala Ser Leu Val Gly Val 275 280 285			920
TCA ACC ACC ACT TTT TTC TAC ATC TTA TGT GGA TGC ATC GGC TAT GCT Ser Thr Thr Thr Phe Phe Tyr Ile Leu Cys Gly Cys Ile Gly Tyr Ala 290 295 300			968
GCA TTT GGA AAC AAT GCC CCT GGA GAT TTC CTC ACA GAT TTC GGG TTT Ala Phe Gly Asn Asn Ala Pro Gly Asp Phe Leu Thr Asp Phe Gly Phe 305 310 315 320			1016
TTC GAG CCC TTT TGG CTC ATT GAC TTT GCA AAC GCT TGC ATC GCT GTC Phe Glu Pro Phe Trp Leu Ile Asp Phe Ala Asn Ala Cys Ile Ala Val 325 330 335			1064
CAC CTT ATT GGT GCC TAT CAG GTG TTC GCG CAG CCG ATA TTC CAG TTT His Leu Ile Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Gln Phe 340 345 350			1112
GAT GAG AAA AAA TGC AAC AGA AAC TAT CCA GAC AAC AAG TTC ATC ACT Val Glu Lys Lys Cys Asn Arg Asn Tyr Pro Asp Asn Lys Phe Ile Thr 355 360 365			1160
TCT GAA TAT TCA GTA AAC GTA CCT TTC CTT GGA AAA TTC AAC ATT AGC Ser Glu Tyr Ser Val Asn Val Pro Phe Leu Gly Lys Phe Asn Ile Ser 370 375 380			1208
CTC TTC AGA TTG GTG TGG AGG ACA GCT TAT GTG GTT ATA ACC ACT GTT Leu Phe Arg Leu Val Trp Arg Thr Ala Tyr Val Val Ile Thr Thr Val 385 390 395 400			1256
GTA GCT ATG ATA TTC CCT TTC TTC AAC GCG ATC TTA GGT CTT ATC GGA Val Ala Met Ile Phe Pro Phe Phe Asn Ala Ile Leu Gly Leu Ile Gly 405 410 415			1304
GCA GCT TCC TTC TGG CCT TTA ACG GTT TAT TTC CCT GTG GAG ATG CAC Ala Ala Ser Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met His 420 425 430			1352
ATT GCA CAA ACC AAG ATT AAG AAG TAC TCT GCT AGA TGG ATT GCG CTG Ile Ala Gln Thr Lys Ile Lys Lys Tyr Ser Ala Arg Trp Ile Ala Leu			1400

435	AAA ACG ATG TGC TAT GTT TGC TTG ATC GTC TCG CTC TTA GCT GCA GCC	440		445		1448
	Lys Thr Met Cys Tyr Val Cys Leu Ile Val Ser Leu Leu Ala Ala Ala					
	450	455		460		
	GGA TCC ATC GCA GGA CTT ATA AGT AGT GTC AAA ACC TAC AAG CCC TTC					1496
	Gly Ser Ile Ala Gly Leu Ile Ser Ser Val Lys Thr Tyr Lys Pro Phe					
	465	470		475	480	
	CGG ACT ATG CAT GAG TGAGTTGAG ATCCTCAAGA GAGTCAAAAA TATATGTAGT					1551
	Arg Thr Met His Glu					
	485					
	AGTTTGGTCT TTCTGTAA CTATCTGGTG TCTAAATCCA ATGAGAATGC TTTATTGCTA					1611
	AAACTTCATG AATCTCTCTG TATCTACATC TTTCAATCTA ATACATATGA GCTCTTCCAA					1671
	AAAAAAAAAA AAAA					1685

, , INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 485 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Lys	Ser	Phe	Asn	Thr	Glu	Gly	His	Asn	His	Ser	Thr	Ala	Glu	Ser
1				5					10					15	
Gly	Asp	Ala	Tyr	Thr	Val	Ser	Asp	Pro	Thr	Lys	Asn	Val	Asp	Glu	Asp
		20					25					30			
Gly	Arg	Glu	Lys	Arg	Thr	Gly	Thr	Trp	Leu	Thr	Ala	Ser	Ala	His	Ile
		35				40					45				
Ile	Thr	Ala	Val	Ile	Gly	Ser	Gly	Val	Leu	Ser	Leu	Ala	Trp	Ala	Ile
	50				55					60					
Ala	Gln	Leu	Gly	Trp	Ile	Ala	Gly	Thr	Ser	Ile	Leu	Leu	Ile	Phe	Ser
65				70				75						80	
Phe	Ile	Thr	Tyr	Phe	Thr	Ser	Thr	Met	Leu	Ala	Asp	Cys	Tyr	Arg	Ala
			85					90						95	
Pro	Asp	Pro	Val	Thr	Gly	Lys	Arg	Asn	Tyr	Thr	Tyr	Met	Asp	Val	Val
		100					105					110			
Arg	Ser	Tyr	Leu	Gly	Gly	Arg	Lys	Val	Gln	Leu	Cys	Gly	Val	Ala	Gln
		115				120					125				
Tyr	Gly	Asn	Leu	Ile	Gly	Val	Thr	Val	Gly	Tyr	Thr	Ile	Thr	Ala	Ser
	130				135					140					
Ile	Ser	Leu	Val	Ala	Val	Gly	Lys	Ser	Asn	Cys	Phe	His	Asp	Lys	Gly

09547404

09854274-051401

145		150		155		160
His Thr Ala Asp Cys Thr Ile Ser Asn Tyr Pro Tyr Met Ala Val Phe						
		165		170		175
Gly Ile Ile Gln Val Ile Leu Ser Gln Ile Pro Asn Phe His Lys Leu						
		180		185		190
Ser Phe Leu Ser Ile Met Ala Ala Val Met Ser Phe Thr Tyr Ala Thr						
		195		200		205
Ile Gly Ile Gly Leu Ala Ile Ala Thr Val Ala Gly Gly Lys Val Gly						
		210		215		220
Lys Thr Ser Met Thr Gly Thr Ala Val Gly Val Asp Val Thr Ala Ala						
		225		230		235
Gln Lys Ile Trp Arg Ser Phe Gln Ala Val Gly Asp Ile Ala Phe Ala						
		245		250		255
r Ala Tyr Ala Thr Val Leu Ile Glu Ile Gln Asp Thr Leu Arg Ser						
		260		265		270
Ser Pro Ala Glu Asn Lys Ala Met Lys Arg Ala Ser Leu Val Gly Val						
		275		280		285
Ser Thr Thr Thr Phe Phe Tyr Ile Leu Cys Gly Cys Ile Gly Tyr Ala						
		290		295		300
Ala Phe Gly Asn Asn Ala Pro Gly Asp Phe Leu Thr Asp Phe Gly Phe						
		305		310		315
Phe Glu Pro Phe Trp Leu Ile Asp Phe Ala Asn Ala Cys Ile Ala Val						
		325		330		335
His Leu Ile Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Gln Phe						
		340		345		350
Val Glu Lys Lys Cys Asn Arg Asn Tyr Pro Asp Asn Lys Phe Ile Thr						
		355		360		365
Ser Glu Tyr Ser Val Asn Val Pro Phe Leu Gly Lys Phe Asn Ile Ser						
		370		375		380
Leu Phe Arg Leu Val Trp Arg Thr Ala Tyr Val Val Ile Thr Thr Val						
		385		390		395
Val Ala Met Ile Phe Pro Phe Phe Asn Ala Ile Leu Gly Leu Ile Gly						
		405		410		415
Ala Ala Ser Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met His						
		420		425		430
Ile Ala Gln Thr Lys Ile Lys Lys Tyr Ser Ala Arg Trp Ile Ala Leu						
		435		440		445
Lys Thr Met Cys Tyr Val Cys Leu Ile Val Ser Leu Leu Ala Ala Ala						
		450		455		460
Gly Ser Ile Ala Gly Leu Ile Ser Ser Val Lys Thr Tyr Lys Pro Phe						

465

470

475

480

Arg Thr Met His Glu
485

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1740 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 80..1558
- (D) OTHER INFORMATION: /product= "amino acid transporter"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTATTTTATA ATTCTCTTC TTTTGTTC TAGCTTTGTA ATTATAGTCT TATTTCTCTT	60
TAAGGCTCAA TAAGAGGAG ATG GGT GAA ACC GCT GCC GCC AAT AAC CAC CGT	112
Met Gly Glu Thr Ala Ala Ala Asn Asn His Arg	
1 5 10	
CAC CAC CAC CAT CAC GGC CAC CAG GTC TTT GAC GTG GCC AGC CAC GAT	160
His His His His His Gly His Gln Val Phe Asp Val Ala Ser His Asp	
15 20 25	
TTC GTC CCT CCA CAA CCG GCT TTT AAA TGC TTC GAT GAT GAT GGC CGC	208
e Val Pro Pro Gln Pro Ala Phe Lys Cys Phe Asp Asp Asp Gly Arg	
30 35 40	
CTC AAA AGA ACT GGG ACT GTT TGG ACC GCG AGC GCT CAT ATA ATA ACT	256
Leu Lys Arg Thr Gly Thr Val Trp Thr Ala Ser Ala His Ile Ile Thr	
45 50 55	
GCG GTT ATC GGA TCC GGC GTT TTG TCA TTG GCG TGG GCG ATT GCA CAG	304
Ala Val Ile Gly Ser Gly Val Leu Ser Leu Ala Trp Ala Ile Ala Gln	
60 65 70 75	
CTC GGA TGG ATC GCT GGC CCT GCT GTG ATG CTA TTG TTC TCT CTT GTT	352
Leu Gly Trp Ile Ala Gly Pro Ala Val Met Leu Leu Phe Ser Leu Val	
80 85 90	
ACT CTT TAC TCC TCC ACA CTT CTT AGC GAC TGC TAC AGA ACC GGC GAT	400
Thr Leu Tyr Ser Ser Thr Leu Leu Ser Asp Cys Tyr Arg Thr Gly Asp	
95 100 105	
GCA GTG TCT GGC AAG AGA AAC TAC ACT TAC ATG GAT GCC GTT CGA TCA	448
Ala Val Ser Gly Lys Arg Asn Tyr Thr Tyr Met Asp Ala Val Arg Ser	
110 115 120	

09854774.051401

ATT CTC GGT GGG TTC AAG	AC AAG ATT TGT GGG TTG AT	AAA TAC TTG	496
Ile Leu Gly Gly Phe Lys	Phe Lys Ile Cys Gly Leu Ile Gln Tyr Leu		
125	130	135	
AAT CTC TTT GGT ATC GCA ATT GGA TAC ACG ATA GCA GCT TCC ATA AGC		544	
Asn Leu Phe Gly Ile Ala Ile Gly Tyr Thr Ile Ala Ala Ser Ile Ser			
140	145 150 155		
ATG ATG GCG ATC AAG AGA TCC AAC TGC TTC CAC AAG AGT GGA GGA AAA		592	
Met Met Ala Ile Lys Arg Ser Asn Cys Phe His Lys Ser Gly Gly Lys			
160	165 170		
GAC CCA TGT CAC ATG TCC AGT AAT CCT TAC ATG ATC GTA TTT GGT GTG		640	
Asp Pro Cys His Met Ser Ser Asn Pro Tyr Met Ile Val Phe Gly Val			
175	180 185		
GCA GAG ATC TTG CTC TCT CAG GTT CCT GAT TTC GAT CAG ATT TGG TGG		688	
Ala Glu Ile Leu Leu Ser Gln Val Pro Asp Phe Asp Gln Ile Trp Trp			
190	195 200		
TC TCC ATT GTT GCA GCT GTT ATG TCC TTC ACT TAC TCT GCC ATT GGT		736	
Leu Ser Ile Val Ala Ala Val Met Ser Phe Thr Tyr Ser Ala Ile Gly			
205	210 215		
CTA GCT CTT GGA ATC GTT CAA GTT GCA GCG AAT GGA GTT TTC AAA GGA		784	
Leu Ala Leu Gly Ile Val Gln Val Ala Ala Asn Gly Val Phe Lys Gly			
220	225 230 235		
AGT CTC ACT GGA ATA AGC ATC GGA ACA GTG ACT CAA ACA CAG AAG ATA		832	
Ser Leu Thr Gly Ile Ser Ile Gly Thr Val Thr Gln Thr Gln Lys Ile			
240	245 250		
TGG AGA ACC TTC CAA GCA CTT GGA GAC ATT GCC TTT GCG TAC TCA TAC		880	
Trp Arg Thr Phe Gln Ala Leu Gly Asp Ile Ala Phe Ala Tyr Ser Tyr			
255	260 265		
TCT GTT GTC CTA ATC GAG ATT CAG GAT ACT GTA AGA TCC CCA CCG GCG		928	
Leu Val Val Leu Ile Glu Ile Gln Asp Thr Val Arg Ser Pro Pro Ala			
270	275 280		
GAA TCG AAA ACG ATG AAG AAA GCA ACA AAA ATC AGT ATT GCC GTC ACA		976	
Glu Ser Lys Thr Met Lys Lys Ala Thr Lys Ile Ser Ile Ala Val Thr			
285	290 295		
ACT ATC TTC TAC ATG CTA TGT GGC TCA ATG GGT TAT GCC GCT TTT GGA		1024	
Thr Ile Phe Tyr Met Leu Cys Gly Ser Met Gly Tyr Ala Ala Phe Gly			
300	305 310 315		
GAT GCA GCA CCG GGA AAC CTC CTC ACC GGT TTT GGA TTC TAC AAC CCG		1072	
Asp Ala Ala Pro Gly Asn Leu Leu Thr Gly Phe Gly Phe Tyr Asn Pro			
320	325 330		
TTT TGG CTC CTT GAC ATA GCT AAC GCC GCC ATT GTT GTC CAC CTC GTT		1120	
Phe Trp Leu Leu Asp Ile Ala Asn Ala Ala Ile Val Val His Leu Val			
335	340 345		
GGA GCT TAC CAA GTC TTT GCT CAG CCC ATC TTT GCC TTT ATT GAA AAA		1168	
Gly Ala Tyr Gln Val Phe Ala Gln Pro Ile Phe Ala Phe Ile Glu Lys			
350	355 360		

09854774.05401

TCA GTC GCA GAG AGA TAT C A GAC AAT GAC TTC CTC AGC AG GAA TTT	1216
Ser Val Ala Glu Arg Tyr Pro Asp Asn Asp Phe Leu Ser Lys Glu Phe	
365 370 375	
GAA ATC AGA ATC CCC GGA TTT AAG TCT CCT TAC AAA GTA AAC GTT TTC	1264
Glu Ile Arg Ile Pro Gly Phe Lys Ser Pro Tyr Lys Val Asn Val Phe	
380 385 390 395	
AGG ATG GTT TAC AGG AGT GGC TTT GTC GTT ACA ACC ACC GTG ATA TCG	1312
Arg Met Val Tyr Arg Ser Gly Phe Val Val Thr Thr Thr Val Ile Ser	
400 405 410	
ATG CTG ATG CCG TTT TTT AAC GAC GTG GTC GGG ATC TTA GGG GCG TTA	1360
Met Leu Met Pro Phe Phe Asn Asp Val Val Gly Ile Leu Gly Ala Leu	
415 420 425	
GGG TTT TGG CCC TTG ACG GTT TAT TTT CCG GTG GAG ATG TAT ATT AAG	1408
Gly Phe Trp Pro Leu Thr Val Tyr Phe Pro Val Glu Met Tyr Ile Lys	
430 435 440	
C AGG AAG GTT GAG AAA TGG AGC ACG AGA TGG GTG TGT TTA CAG ATG	1456
C Arg Lys Val Glu Lys Trp Ser Thr Arg Trp Val Cys Leu Gln Met	
445 450 455	
CTT AGT GTT GCT TGT CTT GTG ATC TCG GTG GTC GCC GGG GTT GGA TCA	1504
Leu Ser Val Ala Cys Leu Val Ile Ser Val Val Ala Gly Val Gly Ser	
460 465 470 475	
ATC GCC GGA GTG ATG CTT GAT CTT AAG GTC TAT AAG CCA TTC AAG TCT	1552
Ile Ala Gly Val Met Leu Asp Leu Lys Val Tyr Lys Pro Phe Lys Ser	
480 485 490	
ACA TAT TGATGATTAT GGACCATGAA CAACAGAGAG AGTTGGTGTG TAAAGTTTAC	1608
Thr Tyr	
CAATTTCAAAG AAAACTCCAA AAATGTGTAT ATTGTATGTT GTTCTCATTT CGTATGGTCT	1668
CATCTTTTGTA ATAAAAATTTA AAACCTTATGT TATAAATTAT AAAAAAAAAA AAAAAAAAAA	1728
AAAAAAAAAA AA	1740

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 493 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Glu Thr Ala Ala Ala Asn Asn His Arg His His His His His
1 5 10 15

Gly His Gln Val Phe Asp Val Ala Ser His Asp Phe Val Pro Pro Gln
20 25 30

005474-05404

Pro Ala Phe Lys Cys Phe . p Asp Asp Gly Arg Leu Lys rg Thr Gly
 35 40 45
 Thr Val Trp Thr Ala Ser Ala His Ile Ile Thr Ala Val Ile Gly Ser
 50 55 60
 Gly Val Leu Ser Leu Ala Trp Ala Ile Ala Gln Leu Gly Trp Ile Ala
 65 70 75 80
 Gly Pro Ala Val Met Leu Leu Phe Ser Leu Val Thr Leu Tyr Ser Ser
 85 90 95
 Thr Leu Leu Ser Asp Cys Tyr Arg Thr Gly Asp Ala Val Ser Gly Lys
 100 105 110
 Arg Asn Tyr Thr Tyr Met Asp Ala Val Arg Ser Ile Leu Gly Gly Phe
 115 120 125
 Lys Phe Lys Ile Cys Gly Leu Ile Gln Tyr Leu Asn Leu Phe Gly Ile
 130 135 140
 . . Ile Gly Tyr Thr Ile Ala Ala Ser Ile Ser Met Met Ala Ile Lys
 145 150 155 160
 Arg Ser Asn Cys Phe His Lys Ser Gly Gly Lys Asp Pro Cys His Met
 165 170 175
 Ser Ser Asn Pro Tyr Met Ile Val Phe Gly Val Ala Glu Ile Leu Leu
 180 185 190
 Ser Gln Val Pro Asp Phe Asp Gln Ile Trp Trp Ile Ser Ile Val Ala
 195 200 205
 Ala Val Met Ser Phe Thr Tyr Ser Ala Ile Gly Leu Ala Leu Gly Ile
 210 215 220
 Val Gln Val Ala Ala Asn Gly Val Phe Lys Gly Ser Leu Thr Gly Ile
 5 230 235 240
 Ser Ile Gly Thr Val Thr Gln Thr Gln Lys Ile Trp Arg Thr Phe Gln
 245 250 255
 Ala Leu Gly Asp Ile Ala Phe Ala Tyr Ser Tyr Ser Val Val Leu Ile
 260 265 270
 Glu Ile Gln Asp Thr Val Arg Ser Pro Pro Ala Glu Ser Lys Thr Met
 275 280 285
 Lys Lys Ala Thr Lys Ile Ser Ile Ala Val Thr Thr Ile Phe Tyr Met
 290 295 300
 Leu Cys Gly Ser Met Gly Tyr Ala Ala Phe Gly Asp Ala Ala Pro Gly
 305 310 315 320
 Asn Leu Leu Thr Gly Phe Gly Phe Tyr Asn Pro Phe Trp Leu Leu Asp
 325 330 335
 Ile Ala Asn Ala Ala Ile Val Val His Leu Val Gly Ala Tyr Gln Val
 340 345 350

09854274-054404